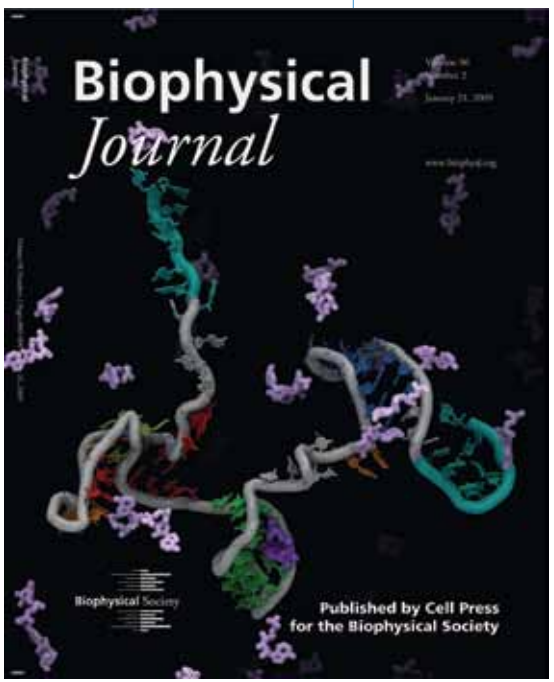
Probing the Dark Matter of the Genome: Mechanistic Studies of Noncoding RNAs

Karissa Y. Sanbonmatsu, Scott P. Hennelly, T-6

Fig. 1. All-atom structure-based simulations performed by Paul Whitford study the interplay between RNA folding and metabolite binding to riboswitches. Gray: S-adenosylmethioine (SAM) riboswitch aptamer domain. Purple: SAM molecules surrounding riboswitch. The simulations revealed that helix P1 (cyan) forms after tertiary core collapse.



It has recently been shown that approximately 98% of the human genome is composed of DNA (deoxyribonucleic acid) that is transcribed into RNA (ribonucleic acid) but does not code for proteins. Many noncoding RNAs have been found to regulate the expression of genes. However the functions of the vast majority of noncoding RNAs have yet to be identified. One noncoding RNA that we can investigate at the molecular level is called the riboswitch. Riboswitches sense the presence of small molecules. If the small molecule is present, a gene is turned off. If the small molecule is not present, a gene is turned on. These molecules were first predicted and synthesized in the lab and later discovered in nature. Approximately 2% of all bacterial genes are regulated by riboswitches. Furthermore, naturally occurring riboswitches have been re-engineered to detect other molecules and to control different aspects of the cell. While the 3D structures of portions of riboswitches have been determined, the mechanism of riboswitch operation is not understood. By understanding this mechanism, we will be able to customize riboswitches to operate in a desired capacity. One such application is a biosensor.

Our overall strategy is to integrate experimentation with simulations at the design level: simulations are used to design experiments and experiments are used to design simulations. For example, all-atom molecular simulations may be used to find positions on the riboswitch more suitable for fluorescent probes for experimentally monitoring conformation. Experiments reveal particular riboswitch positions that are important for metabolite binding. These positions are then studied in detail in the simulations to elucidate a mechanism. In another example, experiments producing melting curves are performed that can be replicated in a simulation. The experimentally determined melting curves then dictate which temperatures to simulate and which force field parameters to use.

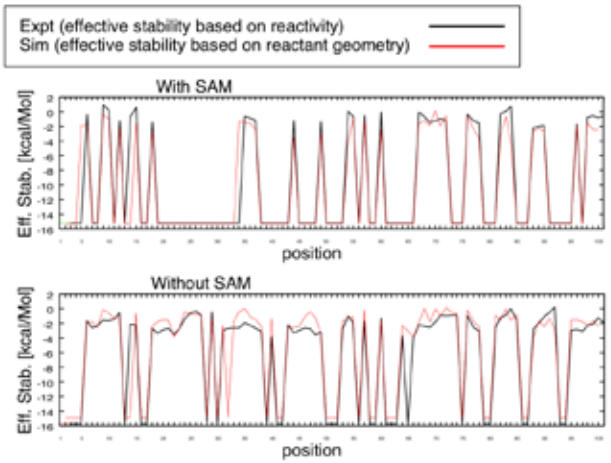


Fig. 2. Comparison between Hennelly's experiments and Hennelly's simulations for equilibrium simulations of SAM riboswitch aptamer. Black: experiment, Red: simulation. Inline probing experiments measure reactivity of RNA backbone cleavage reaction, which indicates relative mobility of backbone. A fitness function based on the geometry of the cleavage reaction is computed from simulations.

As a first step into understanding riboswitch operation, Paul Whitford (T-6) and collaborators at University of California at San Diego performed the first all-atom simulation of a riboswitch aptamer domain (Fig. 1). The aptamer domain is the portion of the riboswitch that binds the small molecule. The expression platform is the portion that turns the gene on or off. This aptamer domain has four helices, P1, P2, P3, and P4. The conventional wisdom in RNA folding states that secondary structure (base pairing and individual helix formation) should form before tertiary structure (contacts between helices). While this sequence of events holds true for most of the helices during our simulation, we find that a functionally important helix, P1, forms after much of the tertiary structure. This occurs because helix

Fig. 3. Caliper liquid-handling robot purchased by Scott Hennelly enables high-throughput inline probing experiments. Reactions are performed on the blue deck. The head unit and gripper (blue and gray block above deck towards right) moves plates around. Each plate allows 96 simultaneous reactions. Twister II arm (right-hand side) moves plates from storage to deck.

P1 requires the tertiary contacts to bring the complementary bases together. This has large implications for how the regulatory decision, on or off, is made.

Our next step was to achieve agreement between simulation and experiment to give us baseline simulation parameters that we can use in future mechanistic studies. Here, Scott Hennelly performed biochemical probing experiments that reveal the secondary structure of the RNA, and also information regarding dynamics of the RNA. The experiments (inline probing and selective 2'-hydroxyl acylation and primer extension (SHAPE probing) monitor reactions along the backbone of the RNA that reflect the mobility of the backbone at each base position along the RNA. For example, nucleotides in helices are shown to be less mobile, while nucleotides in loops and in junctions between helices are more mobile. Hennelly then performed simulations and calculated mobilities of each backbone position. Next a script was written to automatically adjust force-field parameters to achieve better agreement between simulation and experiment. This achieved close agreement between simulation and experiment with a few outlier nucleotides. The parameters for these outlier nucleotides were then individually tuned, resulting in Fig. 2.
To understand the entire riboswitch, including the aptamer domain and the expression platform, Hennelly developed a new experimental assay that monitors switching between the two riboswitch states. By using a two-piece system of RNA instead of only one piece, he is able to observe the switch transitioning from the off state (terminator formed) to the on state (antiterminator formed). Specifically, a fluorescent label molecule is placed on the second strand. If it combines to form the antiterminator, the fluorescence is quenched. Our main result is the discovery that 3D interactions in the sensor domain of the riboswitch are essential for proper riboswitch operation. During normal operation, there is a large difference between switching rates with and without the small molecule. When Hennelly removed 3D interactions by mutating the riboswitch sensor domain, he found that there is no longer a difference. Therefore, the 3D interactions are very important and dependent upon the binding of the small molecule. Hennelly and Sanbonmatsu are both Principal Investigators on a recently funded NIH grant to apply this method to riboswitches.

By performing the first study of the switching mechanism during riboswitch operation, we have laid the groundwork for designing customized biosensors with user-specified switching rates. Often a tradeoff occurs between the sensor domain and the switching domain of the biosensor. We find that 3D interactions dramatically affect this tradeoff. This effect was not known before and will have a large impact on noncoding RNA biosensor design. Designing effective biosensors is the first step to next-generation bioremediation (allowing us to track down hazardous waste) and to more effective chemical agent detectors. The current work not only paves the way for the design of biosensors with novel affinities, but also sensors with low false positives and increased signal. There are also applications in bioenergy, where riboswitches can be used to detect products of cellulose degradation, allowing the engineering of more efficient cellulases (enzymes that break down cellulose).

In the future we plan to use high-throughput platforms and next-generation sequencing methods to design new and more complicated types of RNA switches. Hennelly purchased an automated liquid handling robot (Fig. 3) from LDRD capital equipment funds and is using it for high-throughput secondary structure and RNA dynamics determinations. We plan to use this system, in combination with Illumina sequencing, to perform a large-scale molecular evolution study of RNA switches. In standard molecular evolution, or in vitro selection, RNAs are randomly mutated and more efficient mutants are selected. Successive rounds are performed to optimize the desired characteristics of the RNA. We will use Illumina technology to determine the entire fitness landscape in a single round with much more sampling of sequence space. The promise of riboswitches as either sensors or regulators is in their high specificity and ability to adopt two dramatically different conformations based on binding a small molecule. We have developed a ground-breaking methodology to simultaneously select for these two important characteristics.

For more information contact Karissa Y. Sanbonmatsu at kys@lanl.gov.

Funding Acknowledgments
• LANL Directed Research and Development Program
• NIH, American Recovery & Reinvestment Act (ARRA)